SD211-fiche

Introduction: optimization, machine learning and convex analysis

Classifier : of the kind $h : z \mapsto \text{sign}(\langle x, z\rangle + x_0)$

Elements of convex analysis

Strong convexity : f is μ-stongly convex if $f - \frac{\mu}{2}||.||^2$ is convex.

$\Leftrightarrow f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + \frac{\mu}{2}t(1-t)||x-y||^2$

Lower semi-continuity : cf. p9

convex $\Leftrightarrow f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle$

subdifferential (set of subgradient) : $\partial f(x) = \{\phi : \forall y \in X, f(y) - f(x) \geq \langle \phi, y - x\rangle\}$

Fermat's rule : $x \in \arg\min f \Leftrightarrow 0 \in \partial f(x)$

Primal methods (p.15)

gradient method, **linear search** and **proximal method**

Fenchel-Legendre trandform, dual problem

Fenchel-Legendre conjugate :

$$f^*(\phi) = \sup_{x \in X} \langle phi, x\rangle - f(x)$$

always confex and l.s.c

Fenchel-young : $f(x) + f^*(x) \geq \langle \phi, x\langle$

Fenchel-Moreau : if f convex, l.s.c. and proper, $f = f^{**}$ and $\partial(f^*) = (\partial f)^{-1}$

Lagrangian function : (g inequality constraint, A equality constraint)

$$L : (x, \phi_E, \phi_I) \mapsto f(x) + \langle \phi_E, A(x)\rangle + \langle \phi_I, g(x)\rangle - I_{\mathbb{R}_+^p}(\phi_I)$$

$p = \inf_x \sup_\phi L(x, \phi)$

dual problem : $d = \sup_\phi \inf_x L(x, \phi)$ i.e. maximizing $D(\phi) = \inf_x L(x, \phi)$

Theorem: $d \leq p$

saddle point : minimum over x, maximum over ϕ

Strong duality theorem

Equality constraints

Inequality constraints

Dual methods

Notes over the TDs

$$\text{prox}_{\lambda|.|}(x) = \begin{cases} x - 1 & \text{if } x > \frac{1}{\lambda} \\ 0 & \text{if } x \in [-\frac{1}{\lambda}, \frac{1}{\lambda}] \\ x + 1 & \text{if } x < -\frac{1}{\lambda} \end{cases}$$

$$\text{prox}_{\lambda||.||_2^2}(x) = \frac{1}{1+\lambda}x$$